RELIABILITY OF KILL AND ACTIVITY ESTIMATES IN THE U.S. WATERFOWL HUNTER SURVEY

UNITED STATES DEPARTMENT OF THE INTERIOR
FISH AND WILDLIFE SERVICE
Special Scientific Report — Wildlife No. 240

Library of Congress Cataloging in Publication Data

Couling, Leigh M.
Reliability of kill and activity estimates in the U.S. waterfowl hunter survey.

(Special scientific report—wildlife ; no. 240)
Supt. of Docs. no.: I 49.15/3:240
1. Waterfowl shooting—United States. 2. Hunting surveys—United States. I.
Sen, A. R. II. Martin, Elwood. III. Title. IV. Series.
SK361.A256 no. 240 [SK331] 639.9'79'0973s 81-607087
[338.4'7799244'0973] AACR2

NOTE: Use of trade names does not imply U.S. Government endorsement of commercial products.

RELIABILITY OF KILL AND ACTIVITY ESTIMATES IN THE U.S. WATERFOWL HUNTER SURVEY

By Leigh M. Couling
A. R. Sen
Elwood M. Martin

UNITED STATES DEPARTMENT OF THE INTERIOR
FISH AND WILDLIFE SERVICE
Special Scientific Report—Wildlife No. 240
Washington, D.C. • 1982

Reliability of Kill and Activity Estimates in the U.S. Waterfowl Hunter Survey

by

Leigh M. Couling[1] and A. R. Sen

Canadian Wildlife Service
Migratory Birds Branch
Environment Canada
Ottawa, Ontario, Canada K1A 0E7

and

Elwood M. Martin

Office of Migratory Bird Management
Laurel, Maryland 20708

Abstract

A mail questionnaire survey of waterfowl hunters is conducted each year in the United States to provide information on waterfowl kill and hunter activity. We carried out a study using data from the 1971–72 and 1972–73 hunting seasons to determine the effectiveness of the present U.S. sampling and estimation techniques, and a number of modifications in both sampling and analysis is recommended. We found that stratification of post offices based on the number of duck stamps sold did not give more precise estimates of mean duck bag and days hunted per hunter than estimates obtained in the absence of stratification. Estimation of error variance assuming simple random sampling of hunters instead of cluster sampling of post offices should be avoided because it may lead to significant underestimates. Of the cluster-type estimators examined, the ratio estimator is recommended for estimating means and standard errors of duck bag and days hunted per hunter. Estimates of kill and hunter activity showed wide departures from normality which led to inefficient estimates of the means and of their variances. Log transformation resulted in approximate normality and in considerable increase in precision of the estimates of error of kill. A possible nonresponse problem exists in the stratification scheme because hunters from the larger post offices have lower apparent response rates than those from smaller post offices. Some sampling problems are a result of and dictated by the type of sampling frame currently available; however, it appears that, with the more advanced statistical techniques now available, some improvements can be made in the sampling as well as in the analytical phase.

As part of its continuing efforts toward more efficient management and conservation of wildlife resources, the United States Fish and Wildlife Service (USFWS) initiated a Mail Questionnaire Survey of U.S. waterfowl hunters in 1952. Data obtained in this annual survey are used to estimate U.S. waterfowl hunting activity and success based on a sample of hunters purchasing duck stamps. The reliability of these estimates has been given little consideration in terms of either efficiency or conformity to the sampling scheme (Martin and Carney 1977). The present report will do so by (1) examining the efficiency of the current method of stratification and (2) comparing the efficiency of other estimation techniques with the one presently used. Survey estimates of kill and activity have highly skewed distributions. The effect of this on the estimates and their errors is examined and, where appropriate, methods are recommended for increasing the precision of these estimates. We also briefly examine the important sources of non-sampling error: (1) nonresponse bias due to differences between hunters who do not report their activities and hunters who do, and (2) response bias due to improper or exaggerated reporting of kill.

[1]Present address: S&S Software Ltd., 444 MacLaren St., Ottawa, Ontario, Canada K2P 2C8.

In this study we used hunter kill and activity data from the 1971–72 and 1972–73 season USFWS surveys of 12 of the largest States, including at least one State from each of the four waterfowl flyways that span the continent. The States included were Florida, Maryland, and New York (Atlantic Flyway); Arkansas, Illinois, Louisiana, Minnesota, Missouri, and Wisconsin (Mississippi Flyway); Texas (Central Flyway); and California and Washington (Pacific Flyway).

Design of the U.S. Survey

The U.S. survey sampling frame consists of a master list of post offices that sell duck stamps (Martin and Carney 1977). Each State is stratified into several geographic zones. In each zone, the post offices are further stratified into three groups based on the number of duck stamps sold the previous year. These are stratum 1 (less than 100 stamps), stratum 2 (100 to 1,000 stamps), and stratum 3 (more than 1,000 stamps).

Post offices are selected at random within each stratum. In strata 1 and 2 and in a sub-sample of branch post offices in stratum 3, all duck stamp purchasers are asked to participate in the survey; stratum 3 post offices are usually made up of a number of branch post offices and, to avoid unnecessarily large samples of hunters, a sub-sample of branch offices is chosen. These branch offices are identifiable individually in the sampling phase but are combined under the main post office name in other phases of the analysis because it is not feasible to maintain the identity of each branch office. The post offices are the primary sampling units, and the hunters buying duck stamps are the elements or the ultimate units of selection. Thus in strata 1 and 2, the U.S. survey design is a stratified single-stage cluster sampling scheme, and in stratum 3 there is a second stage consisting of branch post offices. However, for our analysis, it was necessary to treat the large post offices as single units and assume single-stage cluster sampling for all strata.

All post offices selected in the survey are sent name and address forms (survey contact cards) with instructions to give one card to every duck stamp purchaser asking the purchaser to complete the form and return it to the postal clerk for mailing to the USFWS. The purchaser is thus placed on the Service's mailing list. The purchaser retains a portion of the contact card on which an explanation of the survey and a hunting record form for his own use are provided. Actual participation in the survey is variable because some stamp buyers are unwilling to supply their names and addresses and some postal clerks may neglect to hand out cards. Hunters who fill out address cards cannot be viewed as units of a second sampling stage because the decision by hunters on whether to fill out contact cards does not constitute a probability sample.

At the end of the hunting season, each hunter who filled out a contact card is sent a questionnaire to complete and return. Any hunter failing to do so within 3 to 4 weeks is sent a follow-up questionnaire. The data are used to estimate hunting activity and success for each stratum of each zone of each State. The results are then summed to form zone, State, flyway, and national estimates (Martin and Carney 1977).

Estimation of Stratum Means Within a Zone

We will now obtain expressions for estimates of means (and standard errors) for duck kill and number of days hunted per active hunter for a post office stratum within a zone in a State. Because the estimators of the two means have the same mathematical form, we will simply use the term "mean per hunter" in the following discussion.

Let M_i = number of active hunters (those hunting 1 or more days during the season) who purchased duck stamps from the i^{th} post office of a stratum,

y_{ij} = number of ducks killed (or days hunted) by the j^{th} hunter who purchased a duck stamp from the i^{th} post office,

$y_{i.} = \sum_{j=1}^{M_i} y_{ij}$ = number of ducks killed (or days hunted) by all M_i hunters who purchased duck stamps from the i^{th} post office,

n = number of post offices randomly selected from the population of N post offices in a stratum, and

$\bar{\bar{y}} = \sum_{i=1}^{N} y_{i.} / \sum_{i=1}^{N} M_i$ = population mean per hunter for a post office stratum within a zone.

A biased but consistent ratio estimator of the population mean per hunter for a stratum ($\bar{\bar{y}}$) is

$$\hat{\bar{\bar{y}}} = \sum_{i=1}^{n} y_{i.} / \sum_{i=1}^{n} M_i \tag{1}$$

The variance of the estimator is

$$V_{cl}(\hat{\bar{\bar{y}}}) = \frac{N^2(1 - \frac{n}{N})}{n\left(\sum_{i=1}^{N} M_i\right)^2} \left[\frac{\sum_{i=1}^{N} (y_{i.} - M_i \bar{\bar{y}})^2}{N - 1}\right]$$

$$= \frac{1 - \frac{n}{N}}{n\left(\sum_{i=1}^{N} \frac{M_i}{N}\right)^2} \left[\frac{\sum_{i=1}^{N} (y_i - M_i \bar{\bar{y}})^2}{N - 1}\right]. \tag{2}$$

Since the M_i's are known only for the post offices in the sample, their average $\sum_{i=1}^{N} \frac{M_i}{N}$ must be estimated by $\sum_{i=1}^{n} \frac{M_{i.}}{n}$. Hence $V_{cl}(\hat{\bar{\bar{y}}})$ is estimated by

$$v_{cl}(\bar{\bar{y}}) = \frac{n\left(1 - \frac{n}{N}\right)}{\left(\sum_{i=1}^{n} M_i\right)^2} \left[\frac{\sum_{i=1}^{n} (y_{i.} - M_i \bar{\bar{y}})^2}{n - 1} \right]. \quad (3)$$

(The problem of nonresponse in M_i is examined in a later section.)

Appraisal of Post Office Stratification

One reason for stratification of zones into three post office groups was to take advantage of any differences in average hunting activity and success among post offices. It was assumed that hunter characteristics would differ most along a rural–urban axis; stamp sales volume would indicate the rural–urban character. If no such difference existed, this stratification would be inefficient because it would not contribute to the sampling scheme in terms of additional information about hunter characteristics.

We chose two variables, kill per active hunter and days per active hunter, for this examination. One could as well use kill and days per potential hunter, the procedure routinely followed by the USFWS. (A potential hunter is one who bought a duck stamp with the intention of hunting.) Our analysis indicated that the results differ little between the two approaches (because the vast majority of duck stamp buyers go hunting). To determine if stratification was effective, we carried out tests for significant differences between stratum means for both duck kill and days hunted for each of the 12 States (Steel and Torrie 1960).

Although they represent about 36% of all U.S. duck stamp buyers, only about 320 post offices (2% of the sampling frame) sell more than 1,000 duck stamps per season. The entire United States is divided into 188 zones and, although over 50% of the stratum 3 post offices are included in the sample each year, the number sampled in each zone is very small. These differences among strata should be clear from Table 1, which shows the average number of post offices per zone by post office stratum for the States in this study, States which are all much more populous than the average State. In some zones, the number of sample post offices in stratum 2 is also small. Hence, for the 12 States under study, the data for strata 2 and 3 were combined for each zone. The zonal estimates for stratum 1 and for combined strata 2 and 3 were next summed to provide more reliable estimates at the State level, and tests for differences between these totals were then carried out for each State.

The non-normality of kill and days hunted will affect many of the test procedures used for calculating significant differences. Therefore, these normal tests will usually tend to be less powerful than their non-parametric counterparts. However, at State and flyway levels the means will generally be based on fairly large samples to ensure normality and increase the power of test comparisons. Hence, estimates of mean kill and activity are presented at State, flyway, and national levels.

The results of the *t*-tests for differences between stratum 1 and strata 2 and 3 combined for each State (Table 2) show that kill per hunter differed significantly ($P < 0.05$) between strata in only two States each year; mean days hunted differed significantly in three States during 1971–72 and in two States during 1972–73. Thus the number of States to benefit from stratification by post office size was too small to warrant its continued use in the analyses. For other characteristics of interest (e.g., kill per hunter by species, age and sex ratio determination) it may be advantageous to retain this stratification in sampling.

Additional Testing of Zone and State Estimates

Because no differences were found among post office strata, estimates of duck kill per hunter and days hunted per hunter were calculated with and without post office stratification for comparison. Zonal estimates based on post office stratification were obtained as weighted sums of stratum estimates; the weights were proportions of duck stamp sales in each stratum (the method of combining stratum estimates currently being used). Unstratified zone estimates were made by treating all data from a zone as a single sample. State estimates for both methods were obtained as weighted sums of zone estimates, and the weights were proportions of duck stamp sales in each zone. In practice, one would first choose the estimation technique best suited to the State as a whole, apply the method separately to each zone, and then form weighted sums of these results as

$$\hat{\bar{\bar{y}}}_s = \sum_{z=1}^{k} w_z \hat{\bar{\bar{y}}}_z \quad (4)$$

and

$$v(\hat{\bar{\bar{y}}}_s) = \sum_{z=1}^{k} w_z^2 \, v(\hat{\bar{\bar{y}}}_z) \quad (5)$$

where k = number of zones in the State,

$\hat{\bar{\bar{y}}}_z$ = mean per hunter estimate for zone z,

$\hat{\bar{\bar{y}}}_s$ = mean per hunter estimate for State s,

$v(\hat{\bar{\bar{y}}}_z)$ = estimated variance of $\hat{\bar{\bar{y}}}_z$,

$v(\hat{\bar{\bar{y}}}_s)$ = estimated variance of $\hat{\bar{\bar{y}}}_s$, and

$w_z = \dfrac{Q_z}{\sum_{z=1}^{k} Q_z}$ = the weight for zone z where Q_z could be the total number of duck stamp sales in zone z or the number of post offices or unity. When $Q_z = 1$ for all zones, (4) and (5) are of the "unweighted" form.

Table 1. *Average number of post offices available per zone by stratum and State.*

State	1971–72: Average in stratum			1972–73: Average in stratum		
	1	2	3	1	2	3
Florida	31.6	8.7	0.3	28.9	9.7	1.1
Maryland	57.0	15.7	1.1	60.0	20.3	1.7
New York	130.1	18.1	0.6	122.4	35.3	0.7
Arkansas	60.7	19.3	2.3	51.3	29.3	3.3
Illinois	130.6	27.8	1.0	121.0	37.8	1.4
Louisiana	37.2	28.8	3.4	31.8	31.6	4.2
Minnesota	77.0	30.7	3.0	76.3	30.0	3.7
Missouri	76.8	12.2	1.0	68.4	19.6	1.4
Wisconsin	99.4	36.0	3.6	92.2	41.4	5.2
Texas	104.8	23.8	2.2	94.7	29.0	3.7
California	74.7	36.1	4.1	68.3	41.4	5.3
Washington	87.7	29.7	5.0	80.3	28.3	6.7

As expected with proportional sampling, none of the States showed significant differences in means estimated by either method (Table 3). The important finding is that, in most instances, ignoring stratification did not result in decreased precision; on the contrary, there was an apparent gain as indicated by the reduction in the standard errors which may be partly attributable to the small number of post offices on which most of the stratum estimates are based and partly to departures from proportional allocation among strata.

From Cochran (1963:138–139) we estimate gains in efficiency attained when post office stratification is ignored. The term "gain in efficiency" reflects the reduction in the error of estimation obtained when one estimation procedure is replaced by another. Table 4 presents values at the State level obtained by combining zonal estimates for each State. The percentage gains in efficiency of estimates for most States are fairly large. By comparison, the two negative gains are relatively small and might be attributed to other causes such as sampling error. This finding provides further evidence that the present stratification by post office groups within zones is not efficient.

There are reasons, however, to believe that some recognition of the volume of post office duck stamp sales may still be advisable. Mentioned previously is the two-stage sampling often required for the larger post offices involving the use of branch offices, which, under the present accounting system, cannot be treated as separate post offices. There is also the question of biases stemming from lower response rates for larger post offices (discussed in a later section). Finally, because post offices vary so much in size, the number of post offices to be chosen in the sample selection is not a predetermined (constant) value as required by statistical theory; instead, drawing continues until the supply of contact cards available for a stratum is exhausted. Segregating these potential statistical problems into a separate stratum of large post offices may be an appropriate step in isolating and solving them.

Table 2. *Observed probability level (P) for* t*-tests of differences in mean duck kill and days hunted between stratum 1 and strata 2 and 3 combined.*

State	Kill per hunter				Days per hunter			
	1971–72		1972–73		1971–72		1972–73	
	P	df	*P*	df	*P*	df	*P*	df
Florida	0.25[a]	53	0.00	50	0.46	53	0.01	50
Maryland	0.64	55	0.67	49	0.26	55	0.20	49
New York	0.28	112	0.24	106	0.17	112	0.14	106
Arkansas	0.08	34	0.94	28	0.67	34	0.95	28
Illinois	0.01	109	0.88	103	0.09	109	0.69	103
Louisiana	0.00	46	0.00	44	0.41	46	0.94	44
Minnesota	0.30	53	0.69	58	0.64	53	0.92	58
Missouri	0.07	55	0.83	49	0.00	55	0.00	49
Wisconsin	0.06	53	0.74	41	0.03	53	0.70	41
Texas	0.31	79	0.66	80	0.06	79	0.42	80
California	0.30	78	0.44	69	0.02	78	0.26	69
Washington	0.59	34	0.21	38	0.83	34	0.36	38

[a]The observed probability levels are correct to two decimal places.

Table 3. *Comparisons of State level kill-per-hunter and days-per-hunter estimates: stratified by post office size vs. unstratified.*

		Kill per hunter				Days per hunter			
		Stratified		Unstratified		Stratified		Unstratified	
Hunting season and flyway	State	$\bar{x}$	Standard error	$\bar{x}$	Standard error	$\bar{x}$	Standard error	$\bar{x}$	Standard error
1971–72	Florida	7.65	0.83	7.30	0.73	6.20	0.41	6.17	0.35
Atlantic	Maryland	6.02	1.86	6.00	1.73	8.77	0.50	8.83	0.49
	New York	3.91	0.27	3.89	0.26	7.14	0.26	7.09	0.25
Mississippi	Arkansas	9.19	3.14	8.73	2.92	6.70	0.90	6.49	0.84
	Illinois	6.44	0.40	6.20	0.40	8.32	0.49	8.20	0.49
	Louisiana	14.64	1.82	14.37	1.56	8.52	0.72	8.58	0.66
	Minnesota	8.62	0.65	8.54	0.54	7.22	0.44	7.17	0.37
	Missouri	7.68	0.53	7.52	0.51	8.36	0.45	8.15	0.45
	Wisconsin	6.90	0.41	6.63	0.43	8.49	0.57	8.34	0.55
Central	Texas	8.09	0.80	8.29	0.74	6.89	0.38	6.84	0.36
Pacific	California	18.44	1.10	18.31	1.01	8.35	0.34	8.27	0.31
	Washington	9.92	1.00	9.77	0.91	8.87	0.55	8.45	0.54
1972–73	Florida	9.62	0.55	9.35	0.52	6.84	0.39	6.72	0.36
Atlantic	Maryland	5.17	0.83	5.24	0.79	8.71	0.80	8.53	0.78
	New York	4.03	0.32	4.12	0.31	7.04	0.26	7.15	0.25
Mississippi	Arkansas	10.35	1.55	10.28	1.55	8.67	0.92	8.79	0.89
	Illinois	6.97	0.45	6.98	0.45	8.62	0.40	8.60	0.40
	Louisiana	15.07	1.73	14.89	1.52	7.99	0.77	8.04	0.72
	Minnesota	8.78	0.47	8.75	0.48	7.48	0.30	7.46	0.29
	Missouri	6.13	0.52	6.29	0.49	6.84	0.37	6.84	0.36
	Wisconsin	5.79	0.75	5.68	0.69	8.37	0.61	8.35	0.56
Central	Texas	10.70	2.74	11.04	2.58	5.96	0.54	6.05	0.51
Pacific	California	17.01	1.70	17.00	1.61	8.28	0.46	8.28	0.45
	Washington	12.08	1.14	11.82	1.08	8.48	1.03	8.38	0.92

Although the present stratification by post office size has apparently not led to better estimates, stratification by geographic zone should be retained when making State estimates. That is, estimates at the State level should be obtained through some combination of the zonal estimates, because there is much variation between zones in kill per hunter and days per hunter. This significant zone-to-zone variation indicates that stratification of each State into zones was a good addition to the sampling scheme (see Appendix).

Cluster Sampling Versus Simple Random Sampling

It will now be assumed that the data on hunter kill and activity were obtained through simple random selection (SRS) of hunters instead of by cluster sampling as in the current design. This assumption has been made in the limited work done to date by the USFWS for estimating standard errors of the survey estimates, although it is recognized to be only marginally useful for providing a crude approximation. The sample estimate of mean per hunter ($\bar{\bar{y}}$) is given by

$$\hat{\bar{\bar{y}}} = \sum_{i=1}^{n} y_{i.} \Big/ \sum_{i=1}^{n} M_i \qquad (6)$$

which has the same mathematical form as equation (1) though it is based on a different selection scheme.

The variance of this estimate is

$$V_{srs}(\hat{\bar{\bar{y}}}) = \frac{\left(1 - \sum_{i=1}^{n} M_i \Big/ \sum_{i=1}^{N} M_i\right)}{\sum_{i=1}^{n} M_i} \left[\frac{\sum_{i=1}^{N} \sum_{j=1}^{M_i} (y_{ij} - \bar{\bar{y}})^2}{\sum_{i=1}^{N} M_i - 1} \right] \qquad (7)$$

and is estimated by

$$v_{srs}(\hat{\bar{\bar{y}}}) = \frac{\left(1 - \frac{n}{N}\right)}{\sum_{i=1}^{n} M_i} \left[\frac{\sum_{i=1}^{n} \sum_{j=1}^{M_i} (y_{ij} - \hat{\bar{\bar{y}}})^2}{\sum_{i=1}^{n} M_i - 1} \right]. \qquad (8)$$

Table 4. *Gains in the efficiency of State level estimates achieved by ignoring stratification by post office size.*

		Percent gain in efficiency[a]			
		Kill per hunter		Days per hunter	
Flyway	State	1971–72	1972–73	1971–72	1972–73
Atlantic	Florida	29.3	11.9	37.2	17.4
	Maryland	15.6	10.4	4.1	5.2
	New York	7.8	6.6	8.2	8.2
Mississippi	Arkansas	15.6	0.0	14.8	6.9
	Illinois	0.0	0.0	0.0	0.0
	Louisiana	36.1	29.5	19.0	14.4
	Minnesota	44.9	–4.1	41.4	7.0
	Missouri	8.0	12.6	0.0	5.6
	Wisconsin	–9.1	18.1	7.4	18.7
Central	Texas	16.9	12.8	11.4	12.1
Pacific	California	18.6	11.5	20.3	4.5
	Washington	20.8	11.4	3.7	25.3

[a]Gain in efficiency is defined as the relative gain in information where information is the reciprocal of the variance. The gain for estimator 2 (ignoring stratification of a zone by post office size) over estimator 1 (with stratified sampling) is given by

$$G = \frac{\frac{1}{v_2} - \frac{1}{v_1}}{\frac{1}{v_1}} = \frac{v_1 - v_2}{v_2} = \frac{v_1}{v_2} - 1$$

Since the data were obtained through cluster sampling of post offices and not of individuals chosen at random, the above estimate of variance (herein called SRS estimate) will be biased unless the variation among post offices in terms of duck kill or days hunted per hunter is the same as the variation among hunters within post offices, i.e., there is no intracluster correlation (defined below). This relationship is given by the following approximate equation (Cochran 1963:210).

$$V_{cl}(\hat{\bar{\bar{y}}}) \cong \frac{(N\bar{M} - 1)}{\bar{M}(N - 1)} \left[V_{srs}(\hat{\bar{\bar{y}}}) \{1 + (\bar{M} - 1)\rho\} \right] \quad (9)$$

where $\bar{M} = \sum_{i=1}^{n} M_i/n$ and ρ is the intracluster correlation between pairs of hunters buying stamps at the same post office in the cluster sample and is defined as

$$\rho = \frac{E(y_{ij} - \bar{\bar{y}})(y_{iu} - \bar{\bar{y}})}{E(y_{ij} - \bar{\bar{y}})^2}$$

where the numerator is averaged over $\sum_{i=1}^{N} M_i(M_i - 1)/2$ distinct pairs and the denominator over all $\sum_{i=1}^{N} M_i$ values of y_{ij}.

If ρ is greater than zero, then $V_{cl}(\hat{\bar{\bar{y}}})$ will be larger than $V_{srs}(\hat{\bar{\bar{y}}})$. Because $V_{cl}(\hat{\bar{\bar{y}}})$ is an estimate which is appropriate for the cluster sampling scheme, use of $V_{srs}(\hat{\bar{\bar{y}}})$ instead of $V_{cl}(\hat{\bar{\bar{y}}})$ would lead to a biased estimate and should be avoided. Similarly, if ρ is less than zero, $V_{srs}(\hat{\bar{\bar{y}}})$ will overestimate the true variance. However, if $\rho = 0$, then $V_{srs}(\hat{\bar{\bar{y}}})$ and $V_{cl}(\hat{\bar{\bar{y}}})$ are approximately equal (where N is sufficiently large) so that either one can be safely used to estimate the variance. Consequently, the choice of an estimator should be based either on a prior knowledge of the amount of intracluster correlation present in the population or on the amount indicated by data obtained through a sampling of the population. To be reliable, estimates of intraclass correlation should be based on large samples at State, flyway, and national levels.

Results at the State level (Tables 5 and 6) illustrate the relationship in equation (9). For obtaining estimates, the best weighting method (w_z in expressions [4] and [5]) is by duck stamp sales in all post offices (both sampled and unsampled). The results given in Tables 5 and 6 were obtained by using the "unweighted" form because (a) we are comparing two methods, and the weights, if common to both, should have little effect on the comparisons, and (b) some of the tests used become very complex if carried out in other than the unweighted form. This accounts for the slight differences in estimates of kill and days per hunter contained in Tables 5 and 6 as compared with those in Table 3.

Tables 5 and 6 show that for the majority of the States the intracluster correlation coefficient is significantly greater than zero, and in most instances the SRS estimate of standard error is much less than the cluster estimate, an indication that it is underestimating the true error. The percentage underestimate in $V_{srs}(\hat{\bar{\bar{y}}})$ compared with $V_{cl}(\hat{\bar{\bar{y}}})$ is shown in the last columns of the tables. Hence, as suggested by equation (9), estimation must be carried out by the cluster approach.

It must be emphasized that since $V_{srs}(\hat{\bar{\bar{y}}})$ assumes a design other than that which was implemented, its use rather than that of $V_{cl}(\hat{\bar{\bar{y}}})$ is incorrect and should be avoided. Only when the sample intracluster correlation coefficient is zero or close to zero may one utilize $V_{srs}(\hat{\bar{\bar{y}}})$.

In comparing $V_{cl}(\hat{\bar{\bar{y}}})$ with $V_{srs}(\hat{\bar{\bar{y}}})$, N (the number of post offices in the stratum) becomes very important only when it is very small. A small value of N would tend to increase the value of $V_{cl}(\hat{\bar{\bar{y}}})$ with respect to $V_{srs}(\hat{\bar{\bar{y}}})$ which again suggests that $V_{srs}(\hat{\bar{\bar{y}}})$ will underestimate the true variance (or overestimate it depending on whether ρ is positive or negative).

The highly significant and positive intracluster correlation in most of the States for both mean duck kill and days hunted indicates that variation between post offices was greater than variation among hunters within post offices. This result suggests the use of a sampling scheme which selects a large number of post offices but subsamples a small percentage of the hunters from each of these post offices. However, since this is not operationally feasible, a somewhat less efficient procedure would be to obtain a representative sample of post offices by selecting a systematic sample of post offices by location within each zone of each State, and selecting all hunters from the sampled post offices.

Table 5. *Estimated duck kill per hunter, standard error, intracluster correlation, and underestimate in variance obtained at the State, "flyway," and "national" levels by using individual hunters (SRS) rather than post offices (cluster) as sampling units.*

Hunting season and flyway	State	Kill per hunter	Standard error		Intracluster correlation	Underestimate[a] in error (percent)
			Cluster	SRS		
1971–72	Florida	7.15	0.52	0.34	0.0451*[b]	57.2
Atlantic	Maryland	5.71	0.38	0.30	0.2396*	37.7
	New York	4.22	0.33	0.18	0.0457*	70.2
	Combined	5.69	0.26	0.17	0.0838*	57.2
Mississippi	Arkansas	9.47	2.17	0.64	0.1323*	91.3
	Illinois	6.49	0.37	0.23	0.0315*	61.3
	Louisiana	13.48	0.94	0.23	0.0500*	94.0
	Minnesota	8.66	0.32	0.21	0.0263*	56.9
	Missouri	7.74	0.46	0.36	0.0093	—
	Wisconsin	7.20	0.48	0.18	0.0557*	85.9
	Combined	8.79	0.31	0.12	0.0547*	85.0
Central	Texas	7.67	0.36	0.32	0.0325*	21.0
Pacific	California	17.16	0.86	0.61	0.0071	—
	Washington	10.39	0.67	0.51	0.0089	—
	Combined	15.13	0.63	0.45	0.0072	—
Entire Season		8.85	0.19	0.10	0.0170*	72.3
1972–73	Florida	8.80	0.48	0.45	-0.0139	—
Atlantic	Maryland	4.88	0.29	0.28	0.0116	—
	New York	4.52	0.36	0.21	0.0384*	66.0
	Combined	6.35	0.25	0.21	0.0062	—
Mississippi	Arkansas	10.96	0.82	0.77	0.0713*	11.8
	Illinois	7.13	0.43	0.27	0.0330*	60.6
	Louisiana	15.55	0.67	0.27	0.0227*	83.8
	Minnesota	8.80	0.34	0.25	0.0201*	45.9
	Missouri	6.26	0.33	0.29	0.0024	—
	Wisconsin	5.82	0.36	0.20	0.0404*	69.1
	Combined	8.94	0.19	0.13	0.0255*	53.2
Central	Texas	10.30	0.55	0.46	0.0104	—
Pacific	California	16.70	0.73	0.68	0.0204*	13.2
	Washington	12.27	0.80	0.71	0.0417*	23.4
	Combined	15.37	0.56	0.51	0.0246*	17.1
Entire Season		9.39	0.15	0.12	0.0186*	36.0

[a]Underestimate $= \dfrac{[V_{cl}(\hat{\hat{y}}) - V_{srs}(\hat{\hat{y}})]}{V_{cl}(\hat{\hat{y}})} \times 100.$

[b]*Indicates significant correlation with $P < 0.05$.

Estimates for Groups of States

These estimates were obtained in the same manner as was done at the State level. In the following, the terms "flyway" and "national" are used to indicate simulated totals at these levels because only 12 States are represented. For "flyway" level estimates, all zones of all States represented in each flyway (ignoring State boundaries) were combined in "unweighted" form by using equations (4) and (5). To obtain "national" estimates for each hunting season, zone means were combined ignoring both State and flyway boundaries (Tables 5 and 6).

For the majority of the "flyways," the intracluster correlation coefficient is significantly larger than zero and in most of these instances the SRS approach would underestimate the variance by about 72% in the 1971–72 season and 36% in the 1972–73 season for kill per hunter; for days per hunter the figures were respectively 56 and 60%. It is evident that at flyway and national levels, just as at the State level, estimation of variance or error should be made by the cluster technique rather than by SRS. This implies that the zonal estimates must all be calculated by a cluster approach such as equations (1) and (3). State, flyway, and national results are then automatically of the cluster type.

The choice of estimation method can influence subse-

Table 6. *Estimated days hunted per hunter, standard error, intracluster correlation, and underestimate in variance obtained at State, "flyway," and "national" levels by using individual hunters (SRS) rather than post offices (cluster) as sampling units.*

Hunting season and flyway	State	Days per hunter	Standard error		Intracluster correlation	Underestimate[a] in error (percent)
			Cluster	SRS		
1971–72	Florida	5.95	0.29	0.17	0.0541*[b]	65.6
Atlantic	Maryland	8.00	0.29	0.24	0.0018	—
	New York	7.29	0.18	0.15	0.0211*	30.6
	Combined	6.86	0.15	0.10	0.0223*	55.6
Mississippi	Arkansas	7.22	0.59	0.26	0.1316*	80.6
	Illinois	8.38	0.31	0.18	0.0481*	66.3
	Louisiana	8.50	0.30	0.18	0.0604*	64.0
	Minnesota	7.47	0.22	0.16	0.0188*	47.1
	Missouri	8.70	0.39	0.28	0.0451*	48.4
	Wisconsin	8.31	0.35	0.13	0.0667*	86.2
	Combined	8.11	0.14	0.08	0.0567*	67.3
Central	Texas	6.40	0.16	0.14	0.0222*	23.4
Pacific	California	8.50	0.37	0.27	0.0107	—
	Washington	8.82	0.42	0.33	0.0167	—
	Combined	8.60	0.29	0.21	0.0125*	47.6
Entire Season		7.69	0.09	0.06	0.0170*	55.6
1972–73	Florida	6.67	0.22	0.23	0.0028	—
Atlantic	Maryland	7.89	0.29	0.29	0.0082	—
	New York	7.48	0.26	0.19	0.0279*	46.6
	Combined	7.22	0.15	0.13	0.0182*	24.9
Mississippi	Arkansas	9.05	0.64	0.40	0.1052*	60.9
	Illinois	8.80	0.35	0.20	0.0458*	67.3
	Louisiana	8.38	0.25	0.20	0.0487*	36.0
	Minnesota	7.73	0.15	0.16	0.0081	—
	Missouri	6.90	0.35	0.23	−0.0027	—
	Wisconsin	8.27	0.20	0.17	0.0343*	27.7
	Combined	8.10	0.12	0.09	0.0379*	92.4
Central	Texas	6.18	0.23	0.19	0.0110	—
Pacific	California	8.40	0.27	0.31	0.0056	—
	Washington	8.70	0.48	0.35	0.0487*	46.8
	Combined	8.49	0.23	0.24	0.0194*	−8.9
Entire Season		7.72	0.11	0.07	0.0264*	59.5

[a]Underestimate $= \dfrac{[V_{cl}(\bar{\bar{y}}) - V_{srs}(\bar{\bar{y}})]}{V_{cl}(\bar{\bar{y}})} \times 100.$

[b]*Indicates significant correlation with $P < 0.05$.

quent statistical analysis. The results given in Table 5 for the State of Illinois can serve as an example. The hypothesis can be advanced that in Illinois there was no difference between the 1971–72 and 1972–73 hunting seasons in terms of kill per hunter. The estimates of mean duck kill for the two seasons are 6.49 and 7.13, respectively. If the null hypothesis is tested at the 10% significance level by the usual *t*-test, it is accepted when standard error is calculated by the cluster approach but rejected when standard error is calculated by the SRS approach. Because the intracluster correlation coefficients for both seasons are significantly greater than zero, the incorrect choice of the SRS estimator would have led to an unreliable conclusion.

Efficiency of Estimators Based on Cluster Sampling

We have seen that the ratio estimator (equation [1]) based on a simple random sample of post offices (Table 3, ignoring stratification) was most suited for use with data obtained through the U.S. Hunter Survey. Two other estimators will be considered (Cochran 1963:250–252). These are the mean of the unit means (MUM) and the estimate based on probability proportional to an estimate of size (PPES). All three estimators were used to estimate mean duck kill and days hunted per hunter, and coefficients of variation for all 12 States and both hunting seasons (Tables 7 and 8). As before, State estimates are

Table 7. *Estimates of kill per hunter (KPH) and coefficients of variation (C.V.) by using three "cluster" methods.*

Hunting season and flyway	State	Ratio KPH	Ratio 100 C.V.	PPES KPH	PPES 100 C.V.	MUM KPH	MUM 100 C.V.
1971–72	Florida	7.30	7.67	9.03	14.62	7.20	10.56
Atlantic	Maryland	6.00	6.00	10.97	15.86	6.98	17.91
	New York	3.89	5.91	3.95	7.59	3.68	6.25
Mississippi	Arkansas	8.73	12.83	10.61	18.85	19.92	25.65
	Illinois	6.20	5.48	7.65	9.41	5.34	7.68
	Louisiana	14.37	6.12	24.74	15.36	14.53	8.88
	Minnesota	8.54	3.16	15.70	16.81	8.30	4.70
	Missouri	7.52	6.38	9.62	15.59	6.19	7.92
	Wisconsin	6.63	6.64	6.35	8.82	5.72	9.26
Central	Texas	8.29	6.39	10.28	18.29	7.77	11.97
Pacific	California	18.31	3.93	42.58	16.13	18.02	4.72
	Washington	9.77	4.71	17.26	16.63	10.01	8.69
1972–73	Florida	9.35	6.84	7.14	13.45	7.70	7.53
Atlantic	Maryland	5.24	5.92	6.94	19.74	5.12	11.91
	New York	4.12	5.10	5.45	7.71	4.57	7.44
Mississippi	Arkansas	10.28	6.42	15.35	14.14	10.03	12.06
	Illinois	6.98	5.44	13.32	7.06	6.31	6.18
	Louisiana	14.89	3.69	15.30	16.14	15.16	10.55
	Minnesota	8.75	3.09	16.76	17.18	8.50	4.94
	Missouri	6.29	4.45	9.27	17.37	5.90	7.80
	Wisconsin	5.68	4.93	8.02	17.58	6.00	8.00
Central	Texas	11.04	7.79	25.33	28.94	18.89	29.54
Pacific	California	17.00	4.65	29.24	19.32	16.50	7.15
	Washington	11.82	8.63	16.01	16.55	11.14	10.68

weighted sums of zone estimates; the weight for a zone is the proportion of the State's total duck stamp sales that occurred in that zone. Of the three, the estimator most often of highest precision was the ratio estimator, as indicated by the much smaller coefficients of variation generally obtained with this method.

Effect of Non-Normality on the Estimates

Mean ducks shot and days hunted per hunter within each State based on post offices as sampling units were used to estimate departures from normality by skewness

$$(g_1 = m_3/m_2^{3/2}) \text{ and kurtosis } (g_2 = \frac{m_4}{m_2^2} - 3)$$

where

$$m_2 = \sum_i (y_i - \bar{y})^2/n;\ m_3 = \sum_i (y_i - \bar{y})^3/n;$$
$$m_4 = \sum_i (y_i - \bar{y})^4/n;$$

and n is the number of sample post offices in a State. Sample post offices with less than five duck stamp buyers responding were omitted from this study.

The estimates of mean kill, days hunted, skewness, and kurtosis are presented in Tables 9 and 10 (tests of skewness and kurtosis follow Snedecor and Cochran [1967:86–87]). It is evident that kill per hunter was positively and highly skewed in almost all States, which had the effect of increasing the variance of the mean kill and decreasing its precision. Kurtosis was highly pronounced in about 50% of the States, which also reduced the precision of the errors of the means. Both skewness and kurtosis were somewhat lower for mean days hunted than for mean kill.

The question arises as to how large n (the sample size for post offices) must be within a State for the normal approximation to be accurate enough for estimating mean kill and activity. In a personal communication with the second author, W. G. Cochran, whose work (1963:41) assumes only marked skewness in a population, advised us of more recent unpublished work by G. Bartsch for populations in which the deviation from normality involves both marked skewness and kurtosis. Using this information, we have

$$n > 25G_1^2 + 1.64G_2 \qquad (10)$$

where G_1 and G_2 are Fisher's measures of skewness and kurtosis in the population and are estimated by g_1 and g_2, so that a 95% confidence probability statement will not be wrong more than 6% of the time.

Of the 12 States considered in the study, the sample size for estimating kill and activity in 6 (New York, Arkansas, Louisiana, Minnesota, Texas, and Washington), was less

Table 8. *Estimates of days hunted per hunter (DPH) and coefficients of variation (C.V.) by using three "cluster" methods.*

Hunting season and flyway	State	Ratio DPH	Ratio 100 C.V.	PPES DPH	PPES 100 C.V.	MUM DPH	MUM 100 C.V.
1971–72	Florida	6.17	4.70	7.56	12.96	5.84	6.34
Atlantic	Maryland	8.83	3.51	15.80	14.24	8.85	5.65
	New York	7.09	2.26	7.57	6.74	6.73	3.42
Mississippi	Arkansas	6.49	12.33	7.88	13.20	10.06	13.22
	Illinois	8.20	3.41	11.24	8.10	7.71	5.32
	Louisiana	8.58	3.73	15.40	13.12	9.15	6.88
	Minnesota	7.17	2.65	14.33	16.33	7.07	3.11
	Missouri	8.15	4.42	11.82	16.33	7.08	5.93
	Wisconsin	8.34	5.28	10.40	11.15	8.48	6.72
Central	Texas	6.84	3.22	7.54	11.67	6.65	6.77
Pacific	California	8.27	2.66	16.26	12.85	7.74	3.10
	Washington	8.45	4.26	15.37	15.68	8.91	7.63
1972–73	Florida	6.72	3.87	5.98	12.37	6.11	5.73
Atlantic	Maryland	8.53	3.99	10.03	13.16	8.72	8.03
	New York	7.15	2.80	9.45	5.93	7.45	11.79
Mississippi	Arkansas	8.79	6.03	14.80	12.84	9.41	7.86
	Illinois	8.60	3.95	15.45	32.36	7.94	4.66
	Louisiana	8.04	3.11	10.22	18.20	9.00	9.44
	Minnesota	7.46	2.41	15.37	16.07	7.44	2.96
	Missouri	6.84	3.80	11.77	15.72	6.79	5.30
	Wisconsin	8.35	4.79	12.35	15.95	8.49	6.12
Central	Texas	6.05	8.43	11.14	13.64	7.69	11.96
Pacific	California	8.28	2.78	13.13	13.10	7.92	4.29
	Washington	8.38	5.61	12.13	12.61	8.94	7.83

than that needed for application of normal approximation. Because these 12 States had much larger samples than most in the U.S. survey, kill and activity would be estimated in almost all others with even less confidence unless sample sizes were increased considerably, which is not generally feasible.

However, when the data were transformed by setting $x_i = \ln(y_i + 1)$ where $y_i = y_{i.}/M_i$, the mean kill or activity for the i^{th} post office in the sample, the distributions were approximately normal; only 2 of the 12 States showed significant skewness and kurtosis, and this, too, was not consistent over the years. For detecting real differences between States when the distributions were highly skewed, estimates of means on the transformed scale were more precise, as evidenced by the confidence intervals presented in Table 11. There was no significant difference on the original scale in mean kill for 1971 between Minnesota and Missouri or Louisiana and California; however, real differences became evident after adjustment for non-normality. Similarly, for 1972 the adjusted mean kills ($\bar{x}$) were significantly different between Arkansas and Wisconsin and between Arkansas and Illinois, although no differences were shown in means calculated directly ($\bar{y}$). Further improvement in sensitivity of tests can be made by pooling estimates of error on the transformed scale because the transformed data are expected to have a more constant variance.

We will now transform the $\bar{x}$ back into the original variates (i.e., obtain antilogarithms) and see if any gain in efficiency has been achieved. This is important because mean kill (or days hunted) can be readily interpreted and is, therefore, more useful to management than its logarithm. The efficiency of $\bar{y}$ with respect to mean m when the means of the logarithms are transformed back is approximately estimated by

$$\frac{(s^2 + \frac{s^4}{2})}{e^{s^2} - 1} \tag{11}$$

where m is approximately equal to $e^{(\bar{x} + \frac{s^2}{2})} - 1$ and s^2 is an unbiased estimate of the variance of x (Finney 1941); in large samples and for small values of σ^2, the efficiency of $\bar{y}$ as given by (11) can be almost 100%. The efficiency of the direct sample mean kill ranged from 97 to 100% (Table 12); for days hunted, the efficiency was 100% in almost all instances.

The efficiency of direct estimates of population variance with respect to estimates based on the transformed data (Finney 1941:159) is also included in Table 12. This shows that direct estimates of the variance of the mean

Table 9. *Estimates of mean duck kill (K/H), skewness (g_1), and kurtosis (g_2) at the State level.*

State	1971–72 K/H	1971–72 g_1	1971–72 g_2	1972–73 K/H	1972–73 g_1	1972–73 g_2
Florida	7.21 (38)[a]	0.82*[b]	0.07	7.43 (30)	1.05**	0.94
Maryland	5.67 (43)	1.25**	1.90*	5.15 (33)	0.48	-0.32
New York	4.01 (98)	2.07**	3.99**	4.37 (89)	1.91**	5.90**
Arkansas	11.36 (28)	3.30**	12.74**	11.62 (18)	3.01**	8.96**
Illinois	5.82 (76)	1.28**	1.72**	6.95 (72)	1.32**	3.34**
Louisiana	13.78 (39)	1.75**	3.83**	15.20 (34)	1.68**	3.30**
Minnesota	8.63 (50)	3.49**	17.52**	9.14 (50)	0.61*	0.16
Missouri	7.06(40)	0.82*	0.30	5.88 (33)	0.80*	0.87
Wisconsin	5.86 (50)	1.18**	1.28*	5.96 (39)	0.79*	0.23
Texas	7.62 (59)	1.47**	2.73**	12.57 (47)	3.45**	12.34**
California	18.65 (61)	0.85**	1.07*	15.97 (51)	0.76*	0.40
Washington	10.41 (29)	1.38**	2.07*	11.66 (31)	2.06**	4.48**

[a]Figures in parentheses are numbers of post offices on which estimates are based.
[b]*Indicates significance with $P < 0.05$; **indicates significance with $P < 0.01$.

Table 10. *Estimates of mean days hunted (D/H), skewness (g_1), and kurtosis (g_2) at the State level.*

State	1971–72 D/H	1971–72 g_1	1971–72 g_2	1972–73 D/H	1972–73 g_1	1972–73 g_2
Florida	6.54 (38)[a]	−0.32	0.69	6.47 (30)	0.21	−1.19
Maryland	8.72 (43)	1.17**[b]	2.04**	7.81 (33)	1.24**	1.15*
New York	6.97 (98)	1.33**	2.92**	7.44 (89)	0.68**	0.16
Arkansas	8.74 (28)	1.12**	1.60*	9.05 (18)	1.38**	1.90*
Illinois	7.90 (76)	0.67*	−0.09	8.47 (72)	0.22	−0.06
Louisiana	8.78 (39)	1.36**	3.33**	8.56 (34)	0.79*	0.02
Minnesota	7.53 (50)	3.15**	14.81**	7.66 (50)	0.51	0.16
Missouri	7.67 (40)	0.54	0.33	6.60 (33)	1.05**	0.60
Wisconsin	7.53 (50)	1.26**	1.58*	8.17 (39)	0.43	0.49
Texas	6.54 (59)	0.77**	2.49**	6.87 (47)	1.78**	4.66*
California	8.15 (61)	0.56*	−0.02	7.97 (51)	0.35	0.76
Washington	9.17 (29)	1.49**	2.02*	8.66 (31)	1.44**	1.78

[a]Figures in parentheses are numbers of post offices on which estimates are based.
[b]*Indicates significance with $P < 0.05$; **indicates significance with $P < 0.01$.

Table 11. *Comparisons of 95% confidence interval (C.I.) estimates of mean kill calculated from original units* ($\bar{y} = \sum_{i=1}^{n} y_i/n$) *and from transformed data* ($\bar{x} = \sum_{i=1}^{n} x_i/n$ where $x_i = \ln[y_i + 1]$) *transformed back to original units.*

State	95% C.I. for $\bar{y}$ 1971	95% C.I. for $\bar{y}$ 1972	95% C.I. for $\bar{x}$ (Converted to original units) 1971	95% C.I. for $\bar{x}$ (Converted to original units) 1972
Florida	5.71–8.70	6.09–8.77	5.75–8.50	6.62–9.02
Maryland	4.66–6.68	4.26–6.04	5.16–6.89	4.81–6.55
New York	3.41–4.62	3.82–4.92	4.01–4.85	4.44–5.37
Arkansas	7.45–15.27	6.79–16.45	7.46–12.81	8.00–13.74
Illinois	4.99–6.65	5.95–7.94	5.31–6.75	6.05–7.85
Louisiana	10.80–16.70	12.13–18.26	10.49–15.03	12.06–16.95
Minnesota	7.65–9.61	8.18–10.10	8.50–9.97	8.67–10.59
Missouri	5.81–8.31	4.94–6.83	6.05–8.41	5.47–7.39
Wisconsin	4.93–6.79	4.97–6.94	5.42–7.03	5.47–7.32
Texas	6.34–8.90	9.58–15.57	6.36–8.58	10.18–13.46
California	16.34–20.95	13.70–18.23	15.33–20.09	13.07–17.46
Washington	8.72–12.10	9.27–14.05	9.30–12.18	9.78–13.46

Table 12. *Efficiency of direct estimates of sample mean kill and days hunted and their variances relative to estimates obtained after logarithmic transformation.*

	Efficiency of direct mean (%)				Efficiency of direct variance (%)			
	Kill		Days hunted		Kill		Days hunted	
State	1971	1972	1971	1972	1971	1972	1971	1972
Florida	98	99	100	100	51	68	81	84
Maryland	99	99	99	100	62	67	69	77
New York	99	99	100	100	63	67	78	82
Arkansas	97	100	100	99	39	53	71	64
Illinois	99	99	100	100	60	55	78	77
Louisiana	98	99	100	100	54	61	74	76
Minnesota	100	100	100	100	82	76	87	84
Missouri	99	100	99	100	59	69	66	87
Wisconsin	99	99	100	100	64	66	82	80
Texas	98	99	100	100	54	62	81	79
California	99	99	100	100	59	60	81	70
Washington	100	99	100	100	73	67	83	83

kill for a State would generally be inefficient compared with estimates of variance obtained by transforming back the variance of the logarithms; for days hunted, however, direct estimates of the variance of the sample mean were generally of high efficiency.

We suggest, therefore, that direct estimates of mean and variance always be adopted except for estimating standard error of mean kill, which should be obtained by applying logarithmic transformation procedures. There is need for examination of data from other States and for a number of years to obtain confirmatory evidence.

Table 13. *Response rates (%) at the post office stratum level in the 1971–72 and 1972–73 Hunter Questionnaire Surveys for the combined 12-State sample.*

	1971–72			1972–73		
Stratum	Address cards	Question-naires	Total	Address cards	Question-naires	Total
1	44.7	67.6	30.2	35.7	67.0	23.9
2	38.4	66.8	25.6	34.0	66.2	22.5
3	21.2	72.4	15.3	22.4	71.2	15.9
Combined	33.9	68.2	23.1	30.1	67.8	20.4

Nonresponse

We will use the term nonresponse to refer to failure to measure some of the units in the selected sample. Nonresponse at a selected post office or branch office will, therefore, be defined as the ratio of the number of people who fail to return completed questionnaires to the number of potential hunters at the selected outlet (about 1% of stamp purchasers have no intention of hunting but are stamp collectors or wish to support conservation by buying a stamp and thus do not qualify as potential hunters). Although nonresponse does not necessarily affect estimation procedures, it can introduce a serious bias.

Nonresponse in the U.S. mail survey occurs in three stages: (1) the postal clerk may fail to give a card to the hunter; (2) the hunter may choose not to fill out an address card; and (3) after having sent in an address card the hunter may not complete and return the questionnaire. Nonresponse at the first stage may be due to an insufficient supply of cards or the postal clerk failing to hand a card to every duck stamp purchaser. Nonresponse at the second and third stages is dependent on the hunter and may add seriously to the error of the estimates. The hunter's refusal to supply his name and address probably holds more potential for nonresponse bias than his failure to return a questionnaire (Table 13).

The number of contact cards sent to sample post offices is based on an estimate of expected stamp sales during the current season. Table 13 shows that questionnaire response rates are very similar among strata but address card response rates decrease with increases in post office size. The very low response rate for stratum 3 is due to a relatively high nonresponse at the address card level as compared with strata 1 and 2. The percentage of respondents is far too low for reliable estimates if the hunters not responding have activity characteristics which differ from those of respondents and, therefore, may cause a substantial nonresponse bias.

Our estimates of expected duck stamp sales, and therefore of response rates, are poorest for stratum 3, but whether the low response is inherent in the stratification scheme or arises from characteristics of the post offices or of the hunters in stratum 3 is not known. Possibly the large offices and branches of stratum 3 have responsibilities for stamp sales divided among more clerks or the clerks have greater workloads, and these contribute to their large address card nonresponse. Whether this nonresponse is real or an artifact of stratification, it is essential that we determine if it introduces nonresponse

biases in these kill and activity estimates. In the other two strata, nonrespondents may represent a different class of hunters from those who respond (Atwood 1956; Sen 1972; Filion 1974). If there is such a bias, elimination of stratification by post office size will not reduce or control it but simply mask it in the overall sample. In the absence of more information on nonresponse, the present stratification system cannot be utilized to improve the efficiency of our estimates.

Our testing for differences in hunter characteristics among strata was handicapped by the relatively small samples of post offices from stratum 3 and sometimes stratum 2; therefore, we combined them. This weakness of the analysis, and therefore its results, is brought into question again by this finding of substantial differences in hunter response among strata and the increased suspicion that other characteristics may also differ. Only further study can answer these questions and provide appropriate corrective measures where needed. Investigations into methods of inducing increased response both at the address card level and in questionnaire returns appear hopeful. In the United States, a reminder letter sent to sample post offices several weeks before the hunting season opens increases address card response materially and, in the Canadian survey, the questionnaire response rate has been substantially increased by sending a reminder card immediately following the first mailing of the questionnaire. Much more effort must be put into these and related studies if the reliability of our harvest surveys is to meet the increasingly high standards expected of them.

Summary

The present U.S. method of substratifying geographic zones into post office groups based on the number of duck stamps sold is ineffective in terms of estimation of the precision of hunter success and activity figures. Elimination of post office stratification would not decrease the efficiency of the estimates but would decrease the amount of work required for implementation. However, other problems involving post office size and other considerations such as sampling for the closely related duck wing survey might require retaining such stratification. These aspects need further evaluation.

Estimates of error for both ducks killed and days hunted should be consistent with the selection scheme which is essentially one of cluster sampling of post offices and not of individual hunters buying duck stamps at post offices. Estimation by methods other than those based on cluster sampling is not recommended because it can introduce bias.

Intracluster correlation coefficients are almost always positive, which indicates that for the sample data there is more variation in kill and activity among post office clusters than among hunters within post offices. This suggests that a sampling scheme in which fewer hunters are selected, each from a large number of post offices, would be more efficient than one in which the same sample is confined to a few post offices. Because such sampling is not feasible, a systematic sample of post offices by location within zones (i.e., increased emphasis on broader geographic distribution of the sample) with sampling of all hunters in each post office appears to be a more effective approach.

Both kill and days hunted per hunter within a State showed marked skewness from normality and yielded inefficient estimates of means and their errors. Logarithmic transformation of the data resulted in approximate normality and in considerable increase in the precision of the estimated error of mean kill.

Nonresponse was highly pronounced at the "address card" level, especially for large post offices, and might lead to sizable bias in estimates of kill and activity. There is need for investigation into this problem in the interest of overall efficiency and precision.

We extend our thanks to G. E. J. Smith and H. Boyd of the Canadian Wildlife Service and to P. H. Geissler of the U.S. Fish and Wildlife Service for their useful comments and suggestions.

References

Atwood, E. L. 1956. Validity of mail survey data on bagged waterfowl. J. Wildl. Manage. 20(1):1–16.

Cochran, W. G. 1963. Sampling techniques. 2nd Edition. J. Wiley & Sons, Inc., New York. 413 pp.

Filion, F. L. 1974. Estimating bias due to nonresponse in harvest surveys. Can. Wildl. Serv. Biom. Sect. Manuscr. Rep. 5. 57 pp.

Finney, D. J. 1941. On the distribution of a variate whose logarithm is normally distributed. J. R. Stat. Soc., Ser. B., 7:155–161.

Martin, E. M., and S. M. Carney. 1977. Population ecology of the mallard: IV. A review of duck hunting regulations, activity, and success, with special reference to the mallard. U.S. Fish Wildl. Serv., Resour. Publ. 130. 137 pp.

Sen, A. R. 1972. Some nonsampling errors in the Canadian waterfowl mail survey. J. Wildl. Manage. 36(3):951–954.

Snedecor, G. W., and W. G. Cochran. 1967. Statistical methods. 6th Edition. Iowa State University Press, Ames. 593 pp.

Steel, R. G. D., and J. H. Torrie. 1960. Principles and procedures of statistics. McGraw-Hill Co., New York. 481 pp.

Wright, V. L. 1978. Causes and effects of biases on waterfowl harvest estimates. J. Wildl. Manage. 42(2):251–262.

APPENDIX

Stratification of States into geographic zones is an effective procedure (improves accuracy of estimation) if there are significant differences between the zones of a State in terms of hunter characteristics. To test for zone differences within each State, we performed analysis of variance (ANOVA) and the results are presented in the following table.

Observed probability levels (P) *of* F-*values calculated to test for zone differences ("between" vs. "within" zone) for each State.*

Hunting season and flyway	State	$\bar{x}$ duck kill		$\bar{x}$ days hunted	
		P	df	*P*	df
1971–72					
Atlantic	Florida	0.14[a]	6, 48	0.01	6, 48
	Maryland	0.47	2, 55	0.00	2, 55
	New York	0.00	6, 110	0.00	6, 110
Mississippi	Arkansas	0.26	2, 33	0.00	2, 33
	Illinois	0.00	4, 107	0.00	4, 107
	Louisiana	0.01	4, 43	0.12	4, 43
	Minnesota	0.21	6, 48	0.00	6, 48
	Missouri	0.00	4, 52	0.00	4, 52
	Wisconsin	0.00	4, 50	0.01	4, 50
Central	Texas	0.05	5, 79	0.00	5, 79
Pacific	California	0.00	6, 73	0.00	6, 73
	Washington	0.16	2, 34	0.07	2, 34
1972–73					
Atlantic	Florida	0.07	6, 45	0.01	6, 45
	Maryland	0.00	2, 49	0.00	2, 49
	New York	0.00	6, 102	0.00	6, 102
Mississippi	Arkansas	0.14	2, 27	0.27	2, 27
	Illinois	0.15	4, 98	0.01	4, 98
	Louisiana	0.01	4, 41	0.00	4, 41
	Minnesota	0.01	6, 53	0.00	6, 53
	Missouri	0.13	4, 46	0.05	4, 46
	Wisconsin	0.06	4, 38	0.13	4, 38
Central	Texas	0.85	5, 76	0.60	5, 76
Pacific	California	0.01	6, 64	0.00	6, 64
	Washington	0.60	2, 37	0.47	2, 37

[a]The observed probability levels are correct to two decimal places.

As the Nation's principal conservation agency, the Department of the Interior has responsibility for most of our nationally owned public lands and natural resources. This includes fostering the wisest use of our land and water resources, protecting our fish and wildlife, preserving the environmental and cultural values of our national parks and historical places, and providing for the enjoyment of life through outdoor recreation. The Department assesses our energy and mineral resources and works to assure that their development is in the best interests of all our people. The Department also has a major responsibility for American Indian reservation communities and for people who live in island territories under U.S. administration.

GPO 833-744

UNITED STATES
DEPARTMENT OF THE INTERIOR
FISH AND WILDLIFE SERVICE
EDITORIAL OFFICE
AYLESWORTH HALL, CSU
FORT COLLINS, COLORADO 80523

POSTAGE AND FEES PAID
U S DEPARTMENT OF THE INTERIOR
INT 423

NOTE: Mailing lists are computerized. Please return address label with change of address.